

建置女子職業網球比賽之壓力分類與可解釋性人工智慧模型

莊凱程¹、蔡孟勳^{2*}、黃憲鐘³、洪偉欣⁴

¹ 國立中興大學生命科學系

² 國立中興大學資訊管理學系

³ 國立中興大學體育室

⁴ 國立中興大學電機工程學系

摘要

目的：本研究探討女子職業網比賽中的壓力表現，透過機器學習方法針對壓力建置分類模型，用來了解職業比賽過程中的關鍵壓力因素。**方法：**資料來源為 WTA 2024 年度的比賽記錄，資料集內容涵蓋女子職業選手技術統計、破發點、比較時間與網球盤局結果等相關資料。依據總局數、面臨破發點情況與比賽時間來建立壓力分類，分別為高壓力與壓力，使用隨機森林、XGBoost、LightGBM 與 TabNet 等機器學習演算法進行分析，並以 SHAP (Shapley Additive Explanations) 研究各變數對於壓力分類的重要程度。**結果：**XGBoost 演算法的分類準確性高達 98%。SHAP 分析顯示，總局數、破發點、破發面臨次數與比賽時間等指標對於比賽壓力最為關鍵，當比賽呈現高張力拉鋸時，如比賽局數較多或攻守雙方頻繁破發時，機器學習模型結果會將比賽認定為高壓力比賽。**結論：**本研究找出的壓力特徵，可做為職業比賽分析與相關訓練之參考資訊。未來若加入更多與比賽相關的情境變數，如比分趨勢，關鍵賽點資料或選手個人比較資料，能更有效掌握比賽壓力來源。

關鍵詞：女子職業網球、壓力分類、機器學習演算法、壓力特徵

壹、緒論

網球為具高度張力的職業運動，過程涉及競技、選手體能與對戰策略運用，且選手本身也會受到來自勝負壓力的心裡層面影響。在高強度的比賽回或關鍵搶分賽局時，選手常面臨破發點被破、防守壓力或限入比分僵局的各種狀況，這些賽場上即時因素都會影響比較流暢度或決定勝負 (Cowden et al., 2016; Kovalchik, 2016; O'Donoghue, 2002; 蕭丞邑、陳志榮、張怡潔、王俊智, 2022)。過去，相關研究多聚焦於選選手動作分析或選手體能需求，常以問卷調查或觀察法來判斷，方法較為侷限，對於如何量化比賽壓力仍有新興研究的可能性。由於公開資料逐年更新釋出，研究人員可使用更真實的資料來進來進行壓力分析，並以分析結果制訂教練或個人化的比賽應對策略。於職業網球比賽中，面臨破發點、局數拉鋸與長時間比較，常被主觀認為是壓力的關鍵。無論是選手的年齡、個人技巧或是技術，當雙方選手限入比賽拉鋸時，比分落差明顯或面臨破發之情境下，選手之心理負擔常顯著增加，然而受限傳統的分析方法，對於產生這些壓力影響因素過程探討較為片面或單一。近年，隨著機器學習或深度學習技術應用日漸跨足至各產業，生成式人工智慧 (generative artificial intelligence, GenAI)、大語言模型 (large language model, LLM) 也逐步應用於各領域研究，本研究嘗試以資料科學方法來結合運動領域運用，用機器學習來處理大量運動資料，透過資料探勘與特徵選取，從各變數資料中找出影響比賽表現或結果特徵的因素，如預測比賽勝負、比賽球路分析或個人表現等 (Sampaio et al., 2024)。機器學習方法優點在於建立更有彈性的分類架構，適合用於造成選手壓力因素影響的相關研究。然而，若單純以機器學習模型之預測結果來判斷比賽內容，對於教練或選手自身而言，無法有效釐清真正影響表現的因素，因此透過機器學習可解釋性方法則能顯著增加分析成效。SHAP (Shapley Additive Explanations) 為一種解釋量化特徵貢獻的方法 (Lundberg & Lee, 2017)，對研究人員而言，根據 SHAP 結果能更準確辨識經過機器學習模型預測後之各種關鍵因素，並將資料分析結果轉換為對教練或選手作為訓練策略或戰術應用的關鍵因素與可能關聯。

壓力在網球比賽中屬於關鍵因素，在比賽分數接時或爭奪關鍵分數時，選手通常會出現較高的心裡負擔。比賽節奏突然改變、面臨破發點的關鍵賽點與選手自身的抗壓能力，都會影響運動動作流暢度與影響運動技術。壓力會影響選手在比賽回合的判斷力或應對技術，例如高張度比賽時可能導致選手的認知判斷正確率，直接造成選手技術失誤或改比賽節奏。在心裡層面，選手在面對高張力情境會明顯出現較多生理反應，例如：心跳加快、選手注意力分散與肌肉緊繃等生理現象，這些現象都會影響選手競技品質與專注度，因此選手的心裡負擔將會明顯加重，進而影響到後續的回應策略。此外，在比賽過程中面臨的破發點或被破發點、防守與長時間比賽局的情境時，都會是重要的壓力關鍵因素。這些因素與選手情緒或比賽張力高度相關，進而使選手產出不同程度的心裡壓力。

女子職業網球的比賽節奏與擊球技巧有別於男子網球。相較於男子選手，女子選手的發球球速普遍較低，且體格優勢對比賽節奏的掌控影響更為明顯。在女子賽事中，底線抽球回擊是常見的得分方式，因此破發局的成敗往往更具關鍵性。由於較難直接靠發球得分，女子選手通常須透過多次接發球，耐心爭取破發機會，這也使得比賽中壓力情境頻繁出現，且更容易受比賽回合數與局數變化所影響。女子網球比賽中常出現多變的攻防策略，中場節奏轉換也易受當下比賽情境影響。選手在面臨破發點時的應對，往往直接反應出選手在承受壓力下的技術調整能力。整體而言，女子網球賽事中局數多寡對比賽的影響較為顯著，一旦進入比分膠著階段，不僅會加選手之技術與心理負荷，甚至出現放棄比賽的心理狀態。本研究將聚焦於長局數、相互破發點的攻防以及比賽時間長度等因素，深入探討女子網球比賽中的壓力形成。透過數據分析比賽壓力來源，不僅有助於選手更有效理解比賽內容，更能對強化訓練與技術應用帶來實質助益。

近年來，隨著人工智慧技術快速興起，生成式人工智慧與大語言模型技術快速推進，使得資料科學技術更加成熟與廣泛應用於各領域。本研究則嘗試以機器學習演算法來處理大規模的職業運動資料，機器學習方法能在即時多變化的環境下找出影響比賽的關鍵可能因素，如選手技術能力分析、各回合運動特徵或長時間比賽等因素。目前各賽事中，分析人員已開始應用這些方法來試著決定比賽勝負、辨識比賽模式、選手戰術研析等，顯見機器學習演算法在運動科學領域具有發展的潛能 (Sampaio et al., 2024)。機器學習是個黑盒子，訓練出的複雜模型雖然具有相當程度的預測能力，但是對於選手而言，還是無法直觀了解模型要解釋的問題，對於教練或選手而言，輔助效益仍有精進的空間。透過導入可解釋性人工智慧 (explainable AI, XAI) 的方法，如：SHAP (Shapley Additive Explanations) 方法來解釋機器模型的預測結果，指出該指標於對模型預測的貢獻度，幫助理解運動資料分析結果。SHAP 工具可幫助選手快速了解模型的預測指標，如：破發點、比賽局數、比賽時間等因素。SHAP 的每一個點代表一個樣本，顏色表示該關鍵因素的表現量或重要性，紅色為高、藍色為低。透 SHAP 能幫助研究人員與選手快速理解機器模型分析結果的代表意義，例如指出長時間、破發點、比賽局數等可能是對比賽壓力具影響力。這類研究方法也增加機器學習演算法的透明度。

本研究以 WTA 2024 公開比賽資料集，透過機器學方法建立女子網球壓力分類模型，並以 SHAP 呈現壓力分類因素。本研究提出了一套可應用於女子職業網球比賽壓力的人工智慧壓力分類方法。透過使用多種機器學習演算法分別比較壓力分析的準確性，再透過 SHAP 研究哪些關鍵因素對於壓力分析較為顯著與關連程度較高，為教練與選手提供智能化、客觀度與易判讀性的參考資訊，並於未來此應用在各層級賽場，提供教練訓練選手或選手調整自身狀態的參考方案。

貳、研究方法

一、資料來源

本研究資料選自 Jeff Sackmann 整理的 WTA Match Results 公開資料庫，範圍函蓋 2024 年女子職業比賽完整記錄。包括女子選手基本資料、統計技術、破發點、比賽時間與各盤分數變數。因此資料集經過整理且持續更新，非常適合用來探討女子比賽的壓力特徵使用。在研究中如遇資料變數缺漏則會直接排除，避免影響資料可信度，確保分析結果品質。

二、建置壓力指標

本研究為忠實呈現比賽中可能產出壓力的關鍵因素，也參考了常見網球比較的壓力指標，從現有資料找出代表性的變數，並建構壓力分析。壓力分數主要由 4 個面向。包括比賽時間、面臨破發點次數、挽救破發點次數與所有局數等。比賽時間會影響比賽節奏與選手負擔程度，面臨破發次數會增加選手對於勝負預測判的壓力、挽救破發點成功次數則可能增加選手信心、局數越長則會明顯損耗選手體力或影響比賽拉鋸次數。計算上，首先針對每場比賽計算面臨破發點次數與挽救破發點次數，其次，由比賽盤數計算估算總局數是否會出現搶七賽局。最後，將所有相關變數標準化總合為壓力分數。為後續分類方便，用中位數作為切點，把所有比賽分為高壓力比賽與低壓力比賽。

三、資料前處理

在使用機器學習模型處理前，需經過資料清洗與資料標準化。為避免影響模型預測效果，將優先排除缺漏較多的比賽資料，包括極端資料與異常資料檢查，為使各數值變數一致性，本研究使用 Z Score 進行標準化，公式如下 (Moore et al., 2017):

$$Z = \frac{X - \mu}{\sigma}$$

而 X 為原始數值， μ 為該變數的平均數， σ 經標準差處理。透過公式，可減少不同特徵標準造成差異，可有效提升預測模型的穩定。缺漏數值處理，如檢查特徵為空值則補零處理，減少模型缺漏導致訓練異常，確認資料清整完成後，以 80% 與 20% 比例切分為訓練集與測試集，確保模型評估結果具代表性。

四、機器學習與深度學習演算法

本研究使用常見的四種機器學習與深度學習演算法，如隨機森林、XGBoost、LightGBM 與 TabNet (Arik & Pfister, 2021; Breiman, 2001; Chen & Guestrin, 2016; Ke et al., 2017)。這四種機器學習演算法適合處理具備多種變數的資料集，在過去相關運動研究分析都有顯著效果。其中，隨機森林採用多棵決策樹來進行分類，可降低模型的過擬合 (overfitting) 程度。而 XGBoost 與 LightGBM 為上升梯度模型，計算時間較短，同時可處理較為複雜的資料結構。TabNet 則是進階的深度學習演算法，通常用於處理非線

性的表格欄位資料值。本研究將資料集以 80% 與 20% 的比例將資料切分為訓練集與測試集，並以相同的資料來進一步訓練驗證。分類後，以準確性最高的模型再進行機器學習可解釋分析。

本研究使用的機器學習與深度學習模型方法分別說明如下：

1. 隨機森林 (Random Forest)

隨機森林採用多顆的決策樹用以進行分類，不斷的隨機抽樣與選取特徵子集，使模型的 overfitting 程度有效下降。分類最後結果會以各決策樹投票 (voting) 結果決定，相關函數說明如下：

$$\hat{y} = \text{mode}\{h_1(x), h_2(x), \dots, h_T(x)\}$$

其中 $h_t(x)$ 為第 t 顆決策樹的分結果， T 為決策樹的總數。此方法在多變數資料相對穩定與判讀的效果。

2. 極端梯度提升 (eXtreme Gradient Boosting, XGBoost)

XGBoost 是一種梯度提升模型的機器學習方法，用加性方式來持續建立新的樹，並修正前面的預測誤差。模型的相關函數說明如下：

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i)$$

其中 f_k 為第 k 棵樹。模型目標函數包含損失函數與正則化項，說明如下：

$$\mathcal{L} = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k)$$
$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \|w\|^2$$

其中 l 為損失函數， Ω 為正則化項， w 為葉節點權重。此外，XGBoost 也具備高度泛化能力與高效率計算特性。

3. 輕量化梯度提升機器 (Light Gradient Boosting Machine, LightGBM)

LightGBM 和 XGBoost 相似，都是梯度提升模型，但 LightGBM 使用 leaf-wise 的樹生長方法，可以有效的提升模型運算時間與效能，模型相關函數說明如下：

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i)$$

LightGBM 具有單邊梯度取樣 (GOSS) 與 EFB 演算法等特性，用來提升訓練效果，其模型相關函數說明如下：

$$\text{Gain} = \frac{1}{2} \left(\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right) - \gamma$$

G 和 H 變數分別代表一階與二階導數累進數值。這種機器學習法在大量資料運算中能加速訓練的效果。

4. 表格資料注意力式可解釋模型 (TabNet)

TabNet 是近年來 Google 提出來用來處理表格資料的深度學習演算法，使用加性模型注意力機制 (sequential attention) 做為資料為分配，跟其他深度學習演算法差別在於，其模型已具備可解釋性效果。此深度學習演算法運算涵蓋特徵變換、分配注意力與逐步挑選特徵，模型相關函數說明如下：

$$\begin{aligned}z_t &= \text{ReLU}(W_t x) \\a_t &= \text{Sparsemax}(P_t z_t) \\x_{t+1} &= x \odot a_t\end{aligned}$$

其中 z_t 為將特徵轉換後的輸出結果， a_t 為注意力之權重， x_{t+1} 為下一步輸入特徵值， \odot 作為元素相乘使用。TabNet 的資料結構可以清楚呈現模型在不同步驟所注意到的特徵，非常合適用於運動數據的可解釋分析。

5. 模型評估與驗證：

為了評估分類模型的學習效果，本研究採用了加權平均 (Weighted Average) 方法來衡量模型的整體性能。該方法根據每個類別的樣本數量來計算準確率 (Accuracy)、精確率 (Precision)、召回率 (Recall) 及 F1 分數 (F1 Score) 的加權平均值，以反應資料集中不同類別比例對評估結果的影響。這些評估指標是通過混淆矩陣 (Confusion Matrix) 計算得出。在進行三分類的評估過程中，混淆矩陣會針對每個類別單獨計算：當某一類別被視為正類別時，其餘兩個類別則被標記為負類別。混淆矩陣提供了模型在正確和錯誤預測情況下的詳細數據，包括以下指標：真陽性 (True Positives, TP)：模型正確預測為正類別的樣本數。假陽性 (False Positives, FP)：模型錯誤地將負類別預測為正類別的樣本數。真陰性 (True Negatives, TN)：模型正確預測為負類別的樣本數。假陰性 (False Negatives, FN)：模型錯誤地將正類別預測為負類別的樣本數。

	Predicted Positive	Predicted Negative
Actual Positive	TP	FP
Actual Negative	FN	TN

參、結果與討論

一、女子職業比賽資料分析

本研究採用 2024 年 WTA 比賽之完整比賽資料，依據壓力中位數將壓力為別高壓力與低壓力。在進行資料清理之後，總計獲得 2380 場有效比賽資料，而高壓力與低壓力資料比例旗鼓相當，適合本研究後續進行模型的分類需求。本研究結果關鍵技術變

數，包括比賽時間、面臨破發點次數、總比賽數、挽救破發點次數等變數，均明顯呈現分佈差異程度。其中，高壓力在前述指標則有較高的平均值。本研究採用壓力分數的中位數作為壓力分類的切割點，主要考量資料分佈和資料分類穩定。由於比賽壓力分數極有可能因極端比賽的影響導致偏態分布，如果直接採用平均數作為壓力分類切割，分類結果可能不平衡。而中位數對於極端數值較不敏感，為使本研究高壓力與低壓力樣本數均衡平均，對於監督式分類模型在訓練與測試階段穩定性較佳 (Sampaio et al., 2024)。故本研究採用中位數作為壓力分類之切割。在本研究中，我們主要比較不同機器學習演算法在壓力分類的表現，並以中位數作為切割點，雖然未進一步進行多重切割點之敏感性分析，但整體模型呈現一致的分類趨勢，未來將會進一步探討不同切割對於機器學習模型表現的影響。

二、模型分類表現結果

本研究使用隨機森林、XGBoost、LightGBM 與 TabNet 等模型中進行訓練與資料驗證，分別比較在壓力分類的表現差異。圖 1 顯示 Random Forest、XGBoost、LightGBM 與 TabNet 四種模型在比賽壓力分類任務上的整體準確性。研究結果顯示，XGBoost 具有最高的分類準確性 (約 0.98)，LightGBM 次之，而 Random Forest 與 TabNet 的準確性較低。整體而言，梯度提升類模型在本研究情境下展現較佳的分類能力。

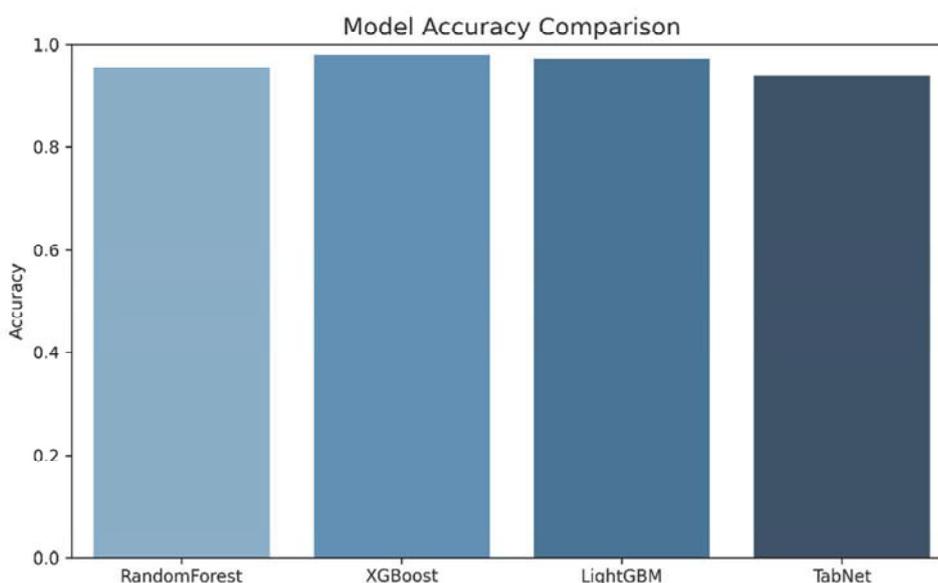


圖 1 四種機器學習模型之分類準確率比較

為進一步比較各模型在不同評估指標上的表現，本研究使用 Precision、Recall 與 F1-score 等指標，並以雷達圖方式呈現整體分類效能。圖 2 以雷達圖呈現四種模型在 Accuracy、Precision、Recall 與 F1-score 四項評估指標上的整體表現。結果顯示，XGBoost 與 LightGBM 在各項指標上皆呈現穩定且接近最高值的分布，顯示其分類效能較為均衡；Random Forest 表現略低但仍具穩定性，而 TabNet 在各指標上的整體表現相對有限。

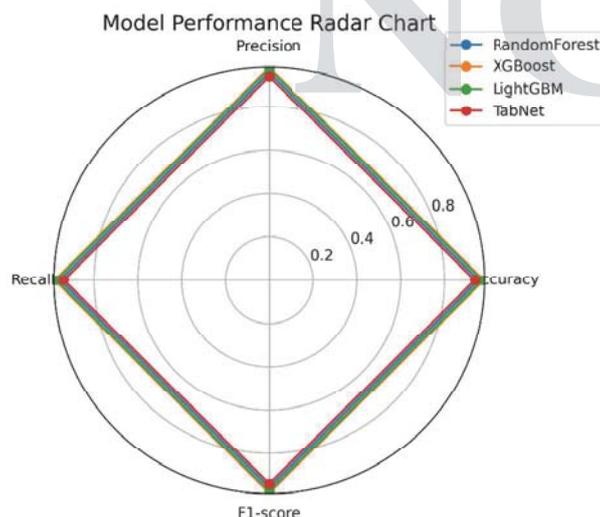


圖 2 各模型分類效能指標之雷達圖

圖 1 與圖 2 結果顯示，XGBoost 的效果最好，本研究進一步將 XGBoost 的結果進行混淆矩陣分析。圖 3 為 XGBoost 模型在測試資料集上的混淆矩陣結果。在測試資料中，高壓力比賽共有 229 場被正確分類，僅 4 場誤判；低壓力比賽則有 238 場正確分類，另有 5 場誤判。由圖 4 可見，高壓力與低壓力比賽皆能被正確分類，多數樣本分布於對角線位置，僅有少數錯誤分類案例。此結果顯示 XGBoost 不僅在整體準確率上表現良好，在實際分類結果上亦具有高度穩定性。

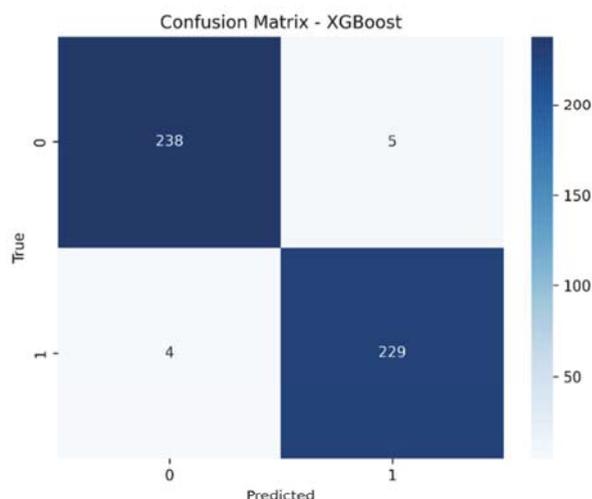


圖 3 XGBoost 模型之混淆矩陣

三、XGBoost 預測結果之 SHAP 與重要特徵分析

圖 1 至圖 3 結果指出，XGBoost 為表現最佳的分類模型。因此，我們進一步結合 SHAP 方法分析影響比賽壓力的關鍵特徵 (Lundberg & Lee, 2017)。透過 SHAP summary 圖，我們可以清楚觀察各變數的影響方法與影響程度。圖四結果顯示，總局數 (total_games) 為最具代表的變數。當比賽局較長時，或不斷拉鋸比賽分數時，SHAP 值

多偏向正向，代表模型在此種比賽情境會傾向判斷為高度壓力比賽，這也與網球比賽觀察一致 (蕭丞邑、陳志榮、張怡潔、王俊智, 2022)，因比賽局數越多代表雙方選手實力不相伯仲，在長時間比賽又無法立即分出勝負就會增加心裡負，明顯拉高比賽張力 (Kovalchik, 2016; O'Donoghue, 2002)。面臨破發點次數 (w_bpFaced、l_bpFaced) 也是模型所關注的重要依據 (Cowden et al., 2016; Sampaio et al., 2024)，SHAP 值散佈明顯，即當選手於比賽過程中持續面臨破發點危機時，SHAP 值往向正向累積，即此變數對於模型預測高度壓力效果顯著。破發點即比賽勝負的關鍵，往往被視為選手於比賽中所面對的高度壓力情境，SHAP 結果進一步印證其關鍵程度。比賽時間 (minutes) 與選手體能高度負荷有關，SHAP 結果顯示長時間比賽也代表更多的比賽回合數與更多的選手體能負荷，因此模型很直觀的判定其為高度壓力情境，此結果相當的符合比賽實際情況，比賽時間長是掌握比賽節奏的重要因素。除了前面所述主要變數外，挽救破發點次數、一發球得分表現 (w_1stWon) 或其他技術指標其實也呈現一定程度的影響，但就 SHAP 而言，分佈過於集中，屬次要貢獻。整體而言，本研究所建構的模型，其分析結果與網球比賽的關鍵變數及實務理解高度吻合。這顯示所採用的機器學習方法，能有效從職業賽事數據中擷取出真實的壓力來源，從而驗證了此模型預測與分析比賽壓力的可行性。

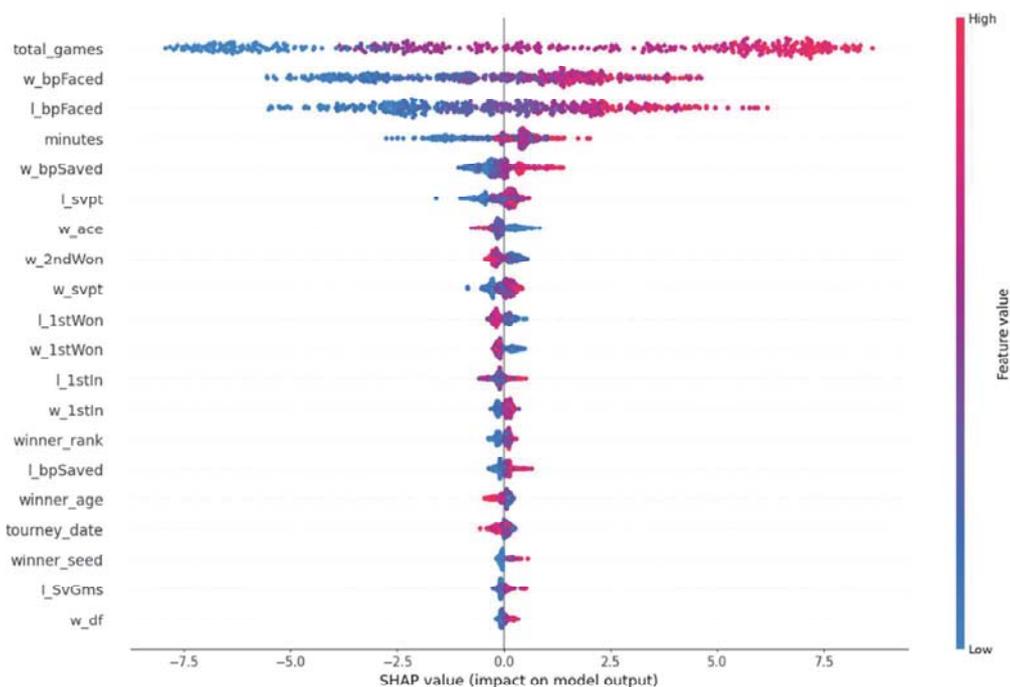


圖 4 XGBoost 模型之 SHAP 特徵重要性分布圖

本研究為進一步比較高壓力與低壓力的差異性，使用統計方法與獨立樣本分析，如表 1。分析結果顯示，高壓力組在總局數 (21.39 ± 7.62)、比賽時間 (119.25 ± 31.38)、以及勝敗雙方破發點面臨次數與挽救次數上，顯著高於低壓力組 ($p < .001$)。故高壓力比賽因普遍伴隨較長的比賽時間與較多的局數拉鋸，且頻繁出現的破發點攻防情境，反映比

賽節的緊湊程度且難分比賽勝負。此研究結果以總局數、破發點相關指標與比賽時間作為壓力量化依據，並與網球比賽高張力比賽情境一致。

表 1
高壓力與低壓力女子職業網球比賽之描述性統計比較

	變數	高壓力組 (Mean ± SD)	低壓力組 (Mean ± SD)	p 值
1	總局數 (total_games)	21.39 ± 7.62	10.05 ± 8.01	<0.001
2	勝方破發點面臨次數 (w_bpFaced)	8.77 ± 4.17	4.41 ± 3.44	<0.001
3	敗方破發點面臨次數 (l_bpFaced)	11.56 ± 3.84	8.09 ± 3.02	<0.001
4	比賽時間 (minutes)	119.25 ± 31.38	92.02 ± 33.64	<0.001
5	勝方破發點挽救次數 (w_bpSaved)	5.40 ± 3.24	2.84 ± 2.53	<0.001
6	敗方破發點挽救次數 (l_bpSaved)	6.08 ± 3.30	3.75 ± 2.51	<0.001

肆、結論

本研究採用公開的 WTA 比賽數據，運用機器學習方法探討女子職業網球比賽中的壓力特徵。結果顯示，高張力賽事與雙方互破發球局的技術指標 (Cowden et al., 2016; O'Donoghue, 2002)，確實可被機器學習模型識別為高壓力比賽。SHAP 分析結果得知，總局數、破發點數量、破發點挽救次數以及比賽時間長度等變數最為關鍵。此發現與過去對網球比賽壓力情境的認知相符，比賽拉鋸時間愈長，選手所累積的壓力愈大，並會直接或間接地影響其臨場表現。本研究採用之四項指標，主要根據網球比賽中常被視為高壓力比賽情境的比賽特徵以機器學習演算法針對壓力定義並以 SHAP 可解釋性找出關鍵特徵 (Lundberg & Lee, 2017)。首先，總局數和比賽時間可反應長時間比賽的拉鋸程度，當比賽局數增加代表雙方選手實力相近，在此情境下，因無法立即分出比賽勝負，因此雙方選的心裡壓力負荷偏高。破發點數量包括勝方破發點面臨次數與敗方破發點面臨次數，在本研究中程式加入雙方之破發點面臨次數，目的是反應比賽中雙方選手遭遇破發危機的情境，也是網球比賽關鍵壓力情境，在破發點面臨時，選手面臨掉分或比賽逆轉的危機，若破發點頻繁出，代表比賽節奏相關緊湊，直接反應選手心裡壓力。破發點挽救次數包括勝方破發點挽救次數與敗方破發點挽救次數，與破發點面臨次數相比，挽救破發點也是反應雙方選手在高壓比賽中情境，也因為造成比賽過程持續延長，並未釋放壓力，只是延後。綜上所述，本研究將總局數、比賽時間、破發點面臨次數與破發點挽救次數等關鍵指標將壓力定義為各種可能情境的綜合結果，以避免用單一指標來判定壓力，也能更全面真實的呈現不同面向的比賽壓力源，使本研究更具參考價值。

從技術層面來看，比賽局數愈多，代表選手需要進行更多回合的抽球回擊與適應比

賽節奏，具備高超技術與高度專注的選手通常表現較佳。當雙方實力相近時，選手很可能調整擊球策略，例如提升底線穩定性、延長對峙時間，或透過觀察落點來改變回球方向。這類即時的技術調整，往往會隨著壓力上升而受到影響，因此總局數成為模型判斷壓力的重要依據，亦符合比賽實際情境。面臨破發點的次數同樣是關鍵因素，破發點被普遍視為高壓情境，選手面對破發點時，除須保持技術穩定性，也承受極大的心理負荷。若比賽中反覆出現破發點的攻防，代表雙方在關鍵分上皆具競爭力，將進一步提升比賽張力。

本研究之 SHAP 分析顯示該變數具有明顯正向影響，代表模型會將破發點頻繁的比賽判定為高壓力賽事，此結果與運動心理學中關於關鍵分心理負荷的論述一致。比賽時間長短也是另一種累積壓力的重要因素，長時間的比賽通常意味著雙方實力接近，選手需運用更多元化的技戰術，包括調動對手、積極跑動，甚至是心理博弈。隨著比賽時間延長，選手體能下降，心理專注度與體力調節也越趨向困難。SHAP 分析該變數呈現正向影響，說明比賽時間本身即為壓力累積的來源。

本研究結果對女子職業網球的訓練與策略擬定具有參考價值，教練可藉由模擬高壓力情境設計訓練，例如增加破發點處理、長時間拉鋸戰或模擬馬拉松式賽局，以強化選手在高壓下的心理素質。然而，本研究仍存在限制，所使用的公開比賽數據未納入選手心理狀態、生理指標與場外因素，因此模型所反應的壓力仍屬客觀賽事指標。未來研究可結合穿戴式裝置與即時心理測量，更全面地解析高強度比賽中的壓力來源。此外，本研究僅分析單一賽季數據，若能納入多個賽季或不同層級賽事資料，將可進一步提升模型之信度與解釋力。

參考文獻

- 蕭丞邑、陳志榮、張怡潔、王俊智 (2022)。心理韌性與心理技能在網球運動之探討。《運動研究》，31(1)，35-45。 [https://doi.org/10.6167/JSR.202206_31\(1\).0004](https://doi.org/10.6167/JSR.202206_31(1).0004)。
- Arik, S. Ö., & Pfister, T. (2021). TabNet: Attentive Interpretable Tabular Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(8), 6679–6687. <https://doi.org/10.1609/aaai.v35i8.16826>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.
- Cowden, R. G., Meyer-Weitz, A., & Oppong Asante, K. (2016). Mental Toughness in Competitive Tennis: Relationships with Resilience and Stress. *Front Psychol*, 7, 320. <https://doi.org/10.3389/fpsyg.2016.00320>
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017).

- LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Neural Information Processing Systems*,
- Kovalchik, S. A. (2016). Searching for the GOAT of tennis win prediction. *Journal of Quantitative Analysis in Sports*, 12(3), 127–138.
- Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Neural Information Processing Systems*,
- Moore, D. S., McCabe, G. P., & Craig, B. A. (2017). *Introduction to the Practice of Statistics*. Macmillan Learning.
- O'Donoghue, P. (2002). Performance models of ladies' and men's singles tennis at the Australian Open. *International Journal of Performance Analysis in Sport*, 2(1), 73–84. <https://doi.org/10.1080/24748668.2002.11868262>
- Sampaio, T., Oliveira, J. P., Marinho, D. A., Neiva, H. P., & Morais, J. E. (2024). Applications of Machine Learning to Optimize Tennis Performance: A Systematic Review. *Applied Sciences-Basel*, 14(13). <https://doi.org/ARTN 551710.3390/app14135517>

Establishing an Artificial Intelligence Model for the Classification and Interpretation of Stress in Women's Professional Tennis

Kai-Cheng Chuang¹, Meng-Hsiun Tsai^{2*}, Hsien-Chung Huang³, Wei-Hsin Hung⁴

¹Department of Life Sciences, National Chung Hsing University, Taichung

²Department of Management Information Systems, National Chung Hsing University,

³Office of physical education and sports, National Chung Hsing University

⁴Department of Electrical Engineering, National Chung Hsing University

Abstract

Purpose: In this study we discussed stress performance in women's professional tennis matches. A machine learning model was developed to classify stress, aiming to understand stress characteristics during professional matches. **Methods:** The data source for this study was WTA 2024 match records, including statistics on female professional players' technical statistics, break points, time and tennis set results. We categorized stress into high stress and stress based on total games played, break-point opportunities faced, and match duration. Machine learning algorithms such as Random Forest, XGBoost, LightGBM, and TabNet were used for stress classification. SHAP (Shapley Additive Explanations) was used to interpret the importance of stress classification in various factors. **Result:** XGBoost achieved the best classification performance, with an accuracy of 98%. The results of SHAP displayed that total games, break points, SHAP identified key variables such as the total number of games, the number of break points faced, and the match duration as crucial factors contributing to match stress. SHAP analysis indicated that total number of games, break points, break-point opportunities faced, and match duration were the most critical indicators of match pressure. When a match exhibits high tension and frequent momentum shifts, the machine learning model classified the match as high-pressure. **Conclusion:** The pressure-related features identified in this study might serve as valuable reference information for professional match analysis and training applications. Future work incorporating additional match-related contextual variables, such as score progression, critical point data, or player-specific comparative information, may further improve the understanding of pressure sources during competition.

Keywords: Women's professional tennis, Stress classification, Machine learning algorithm, Stress characteristics

NCHU